7PAM2000-0105-2021

Applied Data Science 1  
  
  
  
  
Title: Amazon Reviews Mobile Phones

Student id : 21031696  
Name : Narasimha Saphavath

Data science is the field of calculating and processing the scientific and Computer Based algorithm with reference to the various types of structured and unstructured data. This structure and unstructured data are processes to get the data science practices producing various outcomes that can be used in multiple applications and industrial proposes by using the information gained by processing the data

In this study the Amazon mobile phone dataset is used for performing the analysis. The various processes have been performed on the numerical data of the Amazon mobile phones dataset to get the categorical and numerical features of the dataset and found out that the top ten brands as follows

Table

Description automatically generated  
  
 Then extracting the data of the aforementioned brands using the regression and principles of statistics. The correlations are drawn among the various kinds of data that is involved in the below mentioned data set taken from Kaggle to get the relationship between various entities in the dataset and the average of the ratings of the respective brands is shown below

Chart, bar chart

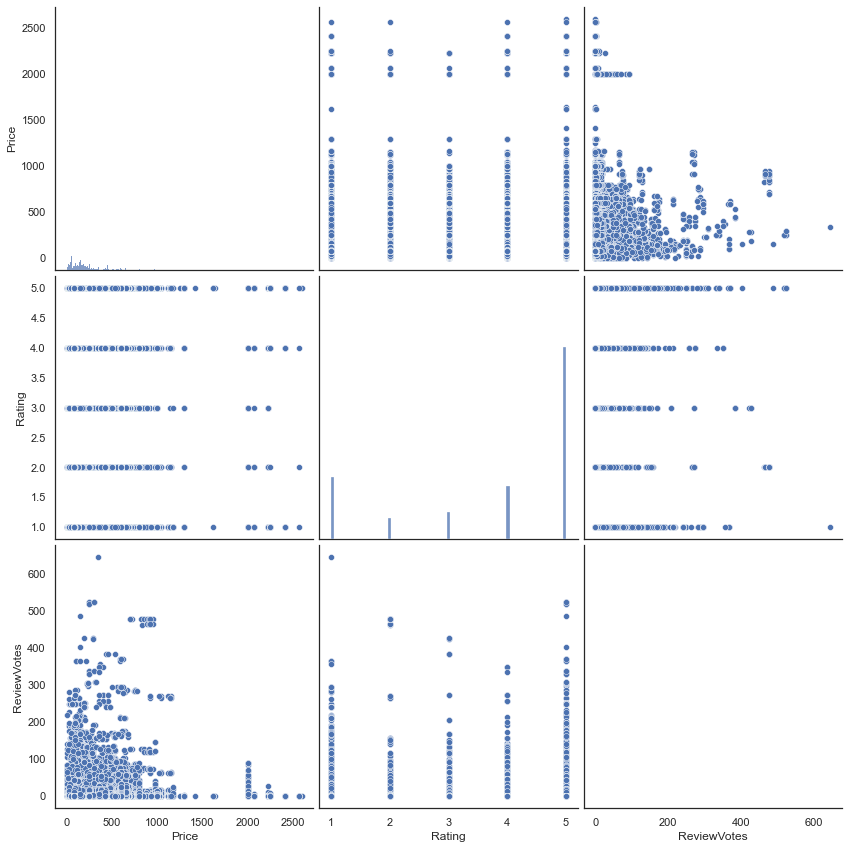
Description automatically generated

The positive correlation is found between the price and rating of the devices

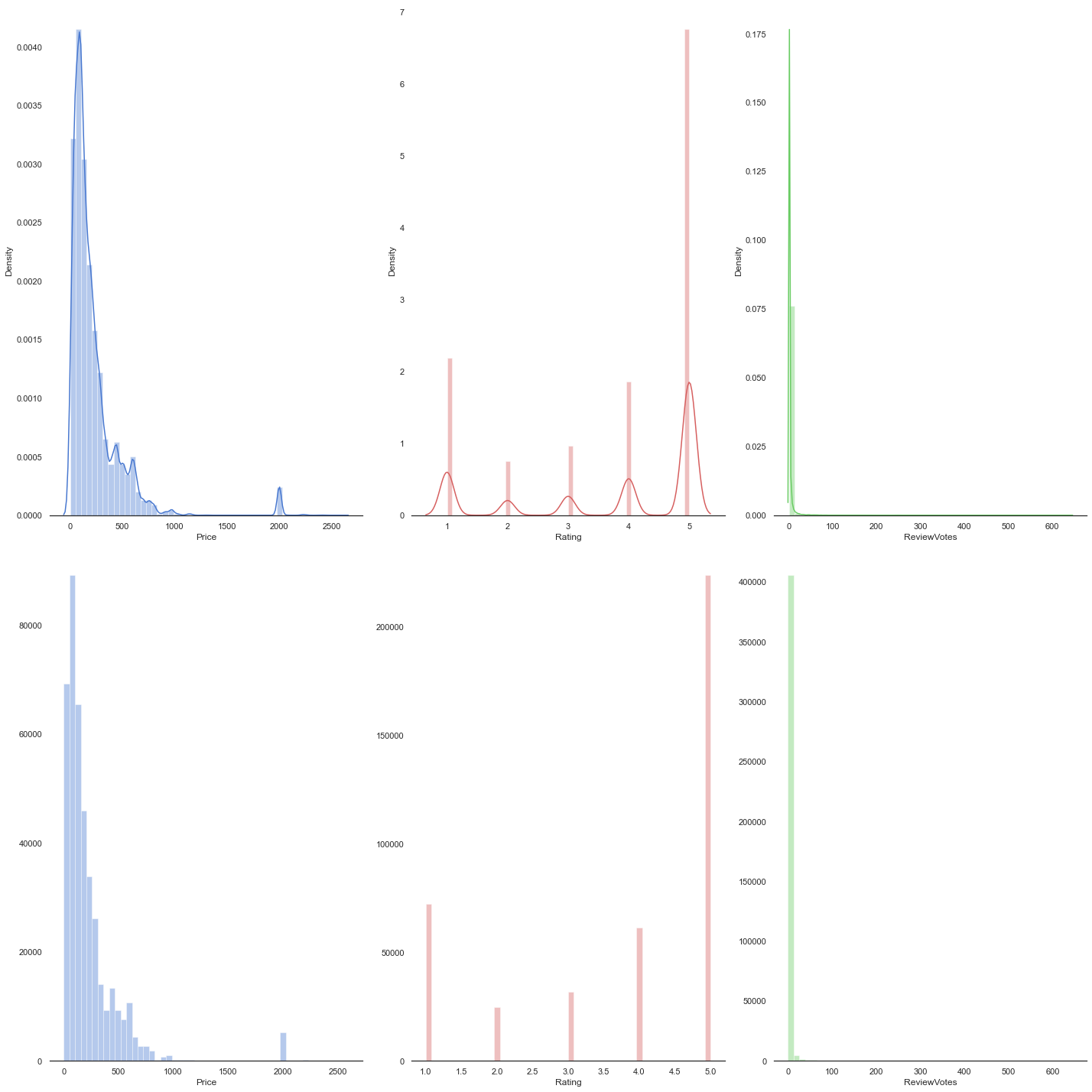
Chart, bar chart

Description automatically generated

The data analysis has been performed on the Amazon mobile phone data set for providing information about the mobile phones. The various brands have been considered along with their review notes, ratings, and some mean statistical information. It should be considerably taking information different users from different brands in the amazon site, the reviews of the products depend majorly over the price and the brand name it is partially shown as complex plot



The data have been expected from the top 10 brands including the discussed brands. By examining data, the most reliable and top brands turn out to be Samsung and Apple. By examining the data, it is tried to find out that the review of the product affects the price related to their brand name. On examining the data, It is seen that the price and review votes are closely related to various user ratings. The rating of the popularity of mobile phones. After analyzing a data, the Samsung is more frequently earning then Apple because of its reviews.



**Conclusion:**

It can be concluded that data Science can be used for producing various information related to the various data sets. It is concluded that Samsung has more popularity and earns more income than Apple because of reviews. It is seen that the most popular brands are Samsung and Apple by examining the data by performing regression analysis.

# **References**

1. Ganorkar, Sandeep S., Nandita Tiwari, and Varsha Namdeo. "Analysis and Prediction of Student Data Using Data Science: A Review." *Smart Trends in Computing and Communications: Proceedings of SmartCom 2020* (2021): 443-448.
2. Hofman, Jake M., Duncan J. Watts, Susan Athey, Filiz Garip, Thomas L. Griffiths, Jon Kleinberg, Helen Margetts et al. "Integrating explanation and prediction in computational social science." *Nature* 595, no. 7866 (2021): 181-188.
3. Sekar, Kalaiarasan, Shahani Aman Shah, A. Antony Athithan, and A. Mukil. "Role of Machine Learning Approaches in Remaining Useful Prediction: A Review." *Intelligent Computing and Innovation on Data Science* (2021): 361-37